



1/21

ESX ORF DNA Sequence (1 to 1116) -> 1-phase Translation  
DNA sequence 1116 b.p. ATGGCTGCAACC ... agtcggaactga linear 371 Amino acids  
MW: 41428 Dalton

1/1  
ATG GCT GCA ACC TGT GAG ATT AGC AAC TAC TTC AGT GCG ATG TAC AGC TCG GAG GAC TCC ACC CTG GCC TCT GTT CCC  
Met ala ala thr cys glu ile ser asn ile phe ser asn tyr phe ser ala met tyr ser glu asp ser thr leu ala ser val pro  
91/31  
CCT GCT GCC ACC TTT GGG GCC GAT GAC TTG GTA CTG ACC CTG AGC AAC CCC CAG ATG TCA TTG GAG GGT ACA GAG AAG GCC AGC TGG TTG  
pro ala ala thr phe gly ala asp asp leu val leu thr leu ser asn pro gln met ser leu glu gly thr glu lys ala ser trp leu  
181/61  
GGG GAA CAG CCC CAG TTC TGG TCG AAG ACG CAG GTT CTG GAC TGG ATC AGC TAC CAA GTG GAG AAG AAC AAG TAC GAC GCA AGC GCC ATT  
gly glu gln pro gln phe trp ser lys thr gln val leu asp trp ile ser tyr gln val glu lys asn lys tyr asp ala ser ala ile  
271/91  
GAC TTC TCA CGA TGT GAC ATG GAT GGC GCC ACC CTC TGC AAT TGT GCC CTT GAG GAG CTG CGT CTG GTC TTT GGG CCT CTG GGG GAC CAA  
asp phe ser arg cys asp met asp gly ala thr leu cys asn cys ala leu glu glu leu arg leu val phe gly pro leu gly asp gln  
361/121  
CTC CAT GCC CAG CTG CAA GAC CTC ACT TCC AGC TCT TCT GAT GAG CTC AGT TGG ATC ATT GAG CTG CTG GAG AAG GAT GGC ATG GCC TTC  
lau his ala gln leu arg asp leu thr ser ser ser asp glu leu ser trp ile ile glu leu leu lys asp gly met ala phe  
451/151  
CAG GAG GCC CTA GAC CCA GGG CCC TTT GAC CAG GGC AGC CCC TTT GCC CAG GAG CTG CTG GAC GAC GGT CAG CAA GCC AGC CCC TAC CAC  
gln glu ala leu asp pro gly pro phe asp gln gly ser pro phe ala gln glu leu leu asp asp gly gln gln ala ser pro tyr his  
541/181  
CCC GGC AGC TGT GGC GCA GGA GCC CCC TCC CCT GGC AGC TCT GAC GTC TCC ACC GCA GGG ACT GGT GCT TCT CCG AGC TCC CAC TCC TCA  
pro gly ser cys gly ala gly ala pro ser pro gly ser ser asp val ser thr ala gly thr gly ala ser arg ser ser his ser ser  
631/211  
GAC TCC GGT GGA AGT GAC GTG GAC CTG GAT CCC ACT GAT GGC AAG CTC TTC CCC AGC GAT GGT TTT CGT GAC TGC AAG AAG GGG GAT CCC  
asp ser gly gly ser asp val asp leu asp pro thr asp gly lys leu phe pro ser asp gly phe arg asp cys lys lys gly asp pro

Fig. 1

811/271	841/281	871/291
AGA GGC ACC CAC CTG TGG GAG TTC ATC CCG GAG CTC AAC GAG GGC CTC ATG AAG TGG GAG AAT CGG CAT GAA		
arg gly thr his leu trp glu phe ile arg asp ile leu ile his pro glu leu asn glu gly leu met lys trp glu asn arg his glu		
901/301	931/311	961/321
GGC GTC TTC aag TTC CTG CGC TCC GAG GCT GTG GCC CAA CTA TGG GGC CAA AAG AAA AAG AAC AGC AAC ATG ACC TAC GAG AAG CTG AGC		
gly val phe lys phe leu arg ser glu ala val ala gln leu trp gly gln lys lys asn ser asn met thr tyr glu lys leu ser		
991/331	1021/341	1051/351
CGG GCC ATG AGG TAC TAC AAA CGG GAG ATC CTG GAA CGG GTG GAT GGC CGG CGA CTC GTC TAC AAG TTT GGC AAA AAC TCA AGC GGC		
arg ala met arg tyr tyr lys arg glu ile leu glu arg val asp gly arg arg leu val tyr lys phe gly lys asn ser ser gly		
1081/361	1111/371	
TGG AAG GAG GAA GAG GTT CTC CAG AGT CGG AAC TGA		
trp lys glu glu glu val leu gln ser arg asn OPA		

Fig. 1 (cont.)

MAATCEISNIFSNYFSAMYSSSEDSTLASVPPAATFGADDLVLTLSNPQMSLEG 53  
 TEKASWLGEGQ**PQFWSKTQVLDNISYQVEKNKYDASAI****DFSRCDMDGATLC****NCA** 106  
 LEEELRVFGPLGDLHAQLRDLTSSSSDELSWIIELLEKDGMAFQEAALDPGPF 159  
 DQGSPEAQELLDDGQQAAPYHPGSCGAG**APSPGSSDVSTAGTGASRSSHSSDS** 212  
**GGSDVDLDP****TDGKLFPSDGF****RDC****KKG****DPKHGKRKRGRPRKLSKEYWDCLEGKK** 265  
 SKHAPRG**T****HLW****EFIR****DL****LI****HP****EL****NE****GL****MK****WENR****HE****GV****EF****K****EL****R****SE****AV****AO****L****W****G****Q****K** 318  
**KKNSNM****TY****E****K****L****S****R****A****M****R****Y****Y****K****R****E****I****L****E****R****V****D****G****R****R****L****V****Y****K****F****G****K****N****S****S****G****W****K****E****E****V****L****Q****S****R****N** 371

Fig. 2A

3/21

Consensus+: P W V W E G LC  
 ESX (64-103): PQFWSKTQVLDWISYQVEKNKYDASAI**DFSRCDMDGATLC**  
 P+ W++T V DW+ + V N++ +DF + M+GA LC  
 ETS-1 (69-106: PRQWTE**TH**VRD**W**MAV--NEFSLKGVD**FQ**KFCMNGAALC

Fig. 2B

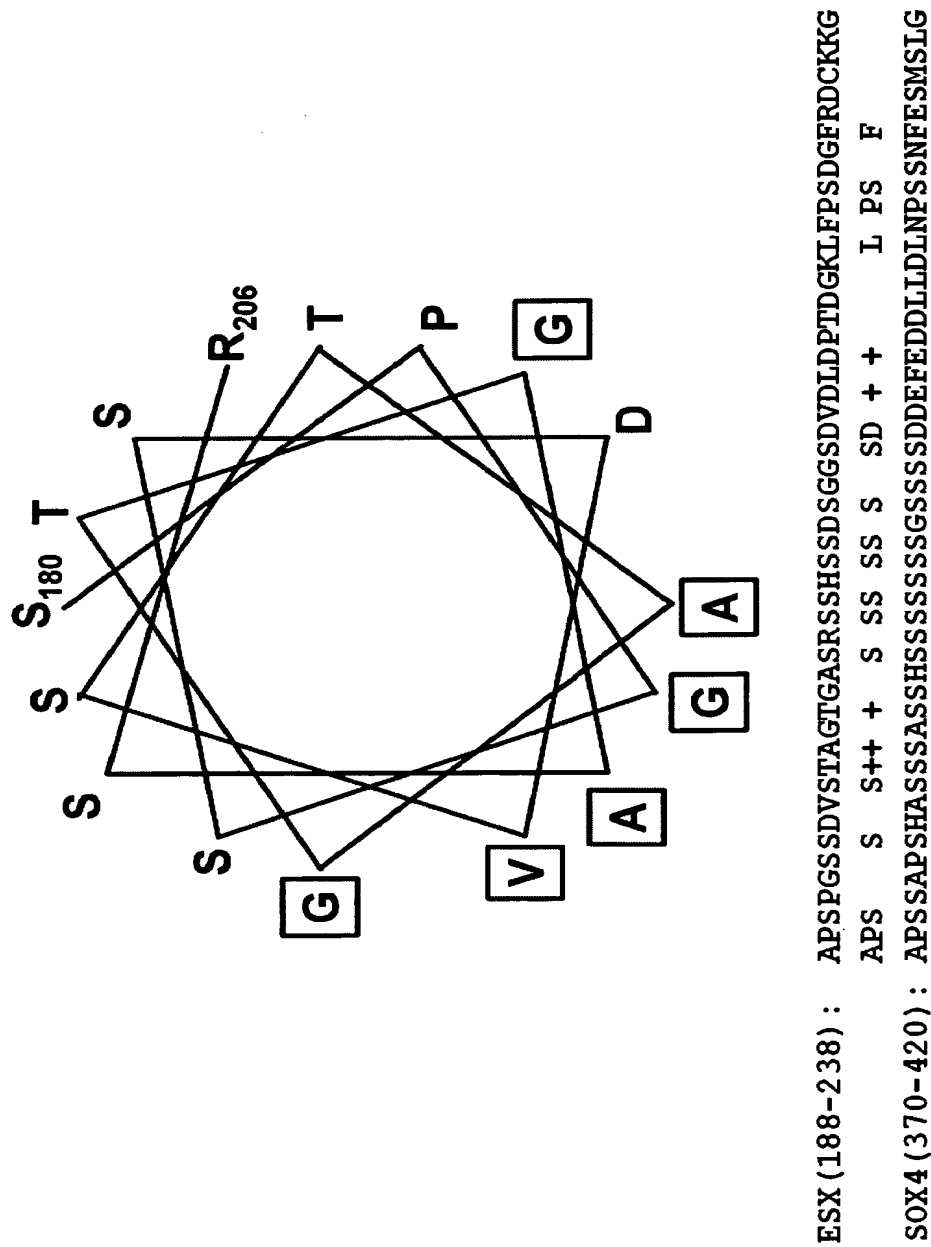


Fig. 2C

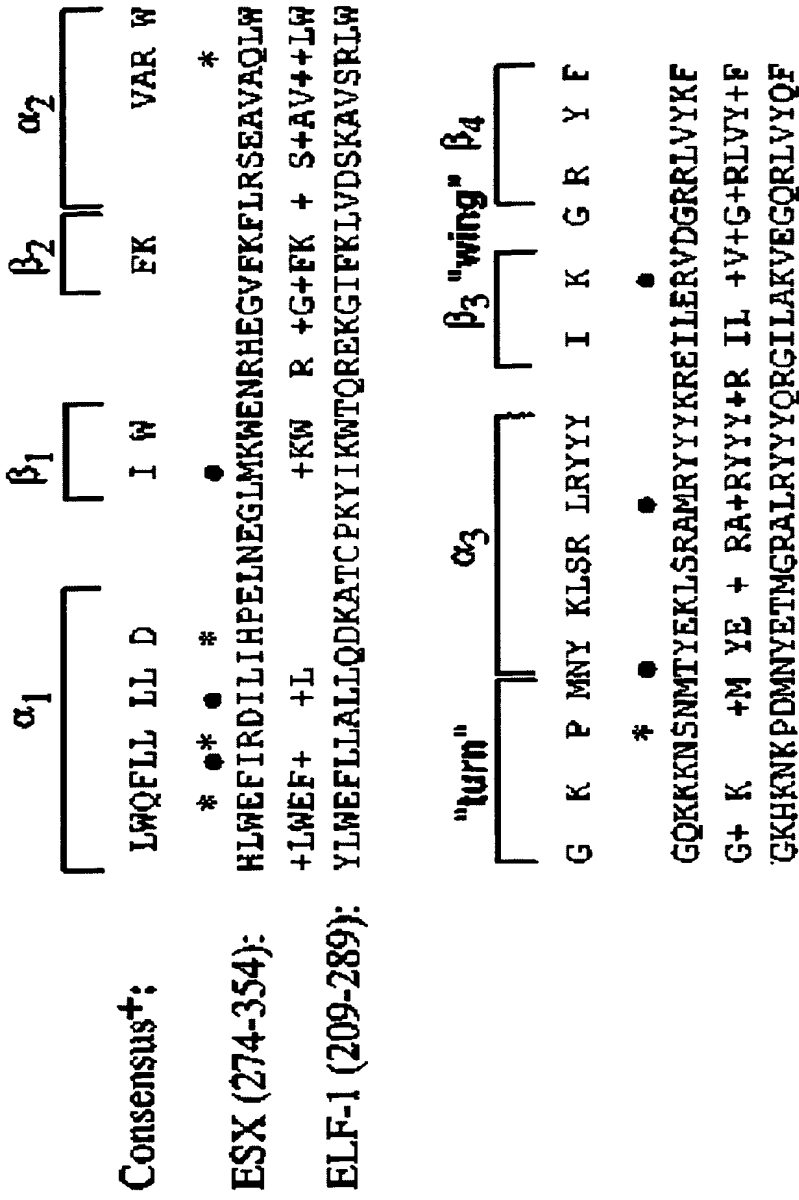


Fig. 2D

## Human ESX Protein Sequence

MAATCEISNIFSNYFSAMYSSDSTLASVPPAATFGADDLVLTLSNPQMSLEG 53  
TEKASWLGEQPQFWSKTQVLDWISYQVEKNKYDASAI DFSRCDMDGATLCNCA 106  
LEELRLVFGPLGDQLHAQLRDLTSSSSDEL SWIIELLEKDGMAFQEA LDPGPF 159  
DQGSPPFAQEL LDDGQQASPYHPGSCGAGAPSPGSSDVSTAGTGASRSSHSSDS 212  
GGSDVDLDPTDGKLFPSDGF RDCKKGD PKHGKRRGRPRKLSKEYWDCLEGKK 265  
SKHAPRGTHLWEFIRD ILIHPELNEGIMKWNRH EGVFKFLRSEAVAQLWGQK 318  
KKNSNMTYEKLSRAMRYYYKREIL ERVDGRRLVYKFGKNSSGWKEE EVLQSRN 371

Fig. 2E

7/21

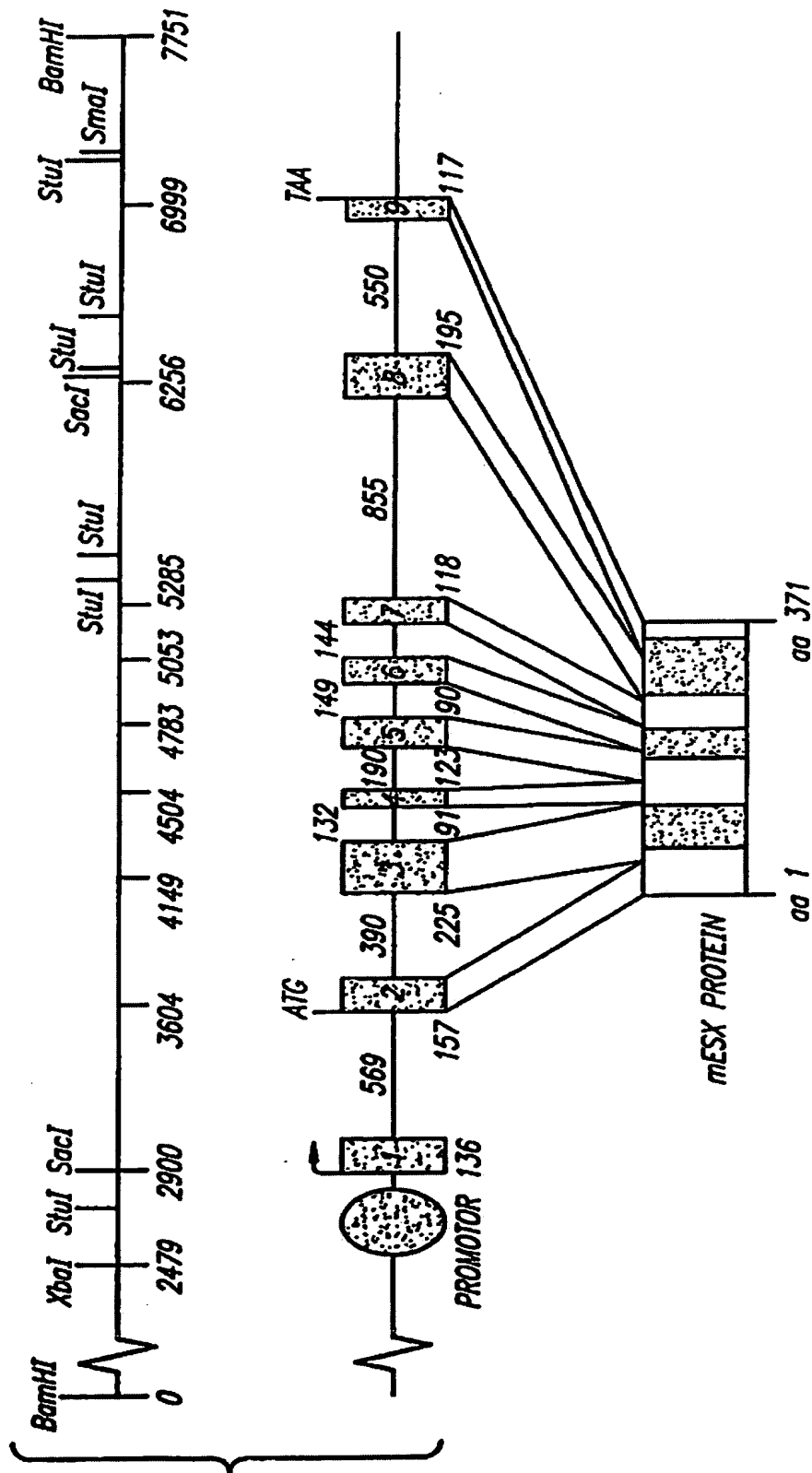


Fig. 3

1/1 31/11  
 ATG GCT GCA ACC TGT GAG ATT AGC AAC ATT TTT AGC AAC TAC TTC AGT GCG ATG TAC AGC  
 Met ala ala thr cys glu ile ser asn ile phe ser asn tyr phe ser ala met tyr ser  
 61/21 91/31  
 TCG GAG GAC TCC ACC CTG GCC TCT GTT CCC CCT GCT GCC ACC TTT GGG GCC GAT GAC TTG  
 ser glu asp ser thr leu ala ser val pro pro ala ala thr phe gly ala asp asp leu  
 121/41 151/51 E2 E3  
 GTA CTG ACC CTG AGC AAC CCC CAG ATG TCA TTG GAG GGT ACA GAG AAC GCC AGC TGG TTG  
 val leu thr leu ser asn pro gln met ser leu glu gly thr glu lys ala ser trp leu  
 181/61 211/71  
 GGG GAA CAG CCC CAG TTC TGG TCG AAG ACG CAG GTT CTG GAC TGG ATC AGC TAC CAA GTG  
 gly glu gln pro gln phe thr trp ser lys thr gln val leu asp trp ile ser tyr gln val  
 241/81 271/91  
 GAG AAG AAC AAG TAC GAC GCA AGC GCC ATT GAC TTC TCA CGA TGT GAC ATG GAT GGC GCC  
 glu lys asn lys tyr asp ala ser ala ile asp phe ser arg cys asp met asp gly ala  
 301/101 331/111  
 ACC CTC TGC AAT TGT GCC CTT GAG GAG CTG CGT CTG GTC TTT GGG CCT CTG GGG GAC CAA  
 thr leu cys asn cys ala leu glu glu leu arg leu val phe gly pro leu gly asp gln  
 361/121 E3 E4 391/131  
 CTC CAT GCC CAG CTG CGA GAC CTC ACT TCC AGC TCT TCT GAT GAG CTC AGT TGG ATC ATT  
 leu his ala gln leu arg asp leu thr ser ser ser asp glu leu ser trp ile ile  
 421/141 451/151 E4 E5  
 GAG CTG CTG GAG AAG GAT GGC ATG GCC TTC CAG GAG GCC CTA GAC CCA GGG CCC TTT GAC  
 glu leu leu glu lys asp gly met ala phe gln glu ala leu asp pro gly pro phe asp

Fig. 4



481/161 511/171  
 CAG GGC AGC CCC TTT GCC CAG GAG CTG CTG GAC GAC GGT CAG CAA GCC AGC CCC TAC CAC  
 gln gly ser pro phe ala gln glu leu leu asp asp gly gln ala ser pro tyr his  
 541/181 571/191 E5 E6  
 CCC GGC AGC TGT GGC GCA GGA GCC Ccc TCC CCT GGC AGC TCT GAC GTC TCC ACC GCA GGG  
 pro gly ser cys gly ala gly ala pro ser pro gly ser ser asp val ser thr ala gly  
 601/201 631/211  
 ACT GGT GCT TCT CGG AGC TCC CAC TCC TCA GAC TCC GGT GGA AGT GAC GTG GAC CTG GAT  
 thr gly ala ser arg ser ser his ser ser asp ser gly gly ser asp val asp leu asp  
 661/221 E6 E7 E8 E9  
 CCC ACT GAT GGC AAG CTC TTC CCC AGC GAT GGT TTT CGT GAC TGC AAG AAG GGG GAT CCC  
 pro thr asp gly lys leu phe pro ser asp gly phe arg asp cys lys lys gly asp pro  
 721/241 751/251  
 AAG CAC GGC AAG CGG AAA CGA GGC CGG CCC CGA AAG CTG AGC AAA GAG TAC TGG GAC TGT  
 lys his gly lys arg lys arg gly arg pro arg lys leu ser lys glu tyr trp asp cys  
 781/261 E7 E8  
 CTC GAG GGC AAG AAG AGC AAG CAC GCG CCC AGA GGC ACC CAC CTG TGG GAG TTC ATC CGG  
 leu glu gly lys ser lys his ala pro arg gly thr his leu trp glu phe ile arg  
 841/281 871/291  
 GAC ATC CTC ATC CAC CCG GAG CTC AAC GAG GGC CTC ATG AAG TGG GAG AAT CGG CAT GAA  
 asp ile leu ile his pro glu leu asn glu gly leu met lys trp glu asn arg his glu  
 901/301 931/311  
 GGC GTC TTC aag TTC CTG CGC TCC GAG GCT GTG GCC CAA CTA TGG GGC CAA AAG AAA AAG  
 gly val phe lys phe leu arg ser glu ala val ala gln leu trp gly gln lys lys lys

Fig. 4 (cont.)

10/21

961/321  
AAC AGC AAC ATG ACC TAC GAG AAG CTG AGC CGG GCC ATG AGG TAC TAC AAA CGG GAG  
asn ser asn met thr tyr glü lys leu ser arg ala met arg tyr tyr lys arg glü  
1021/341  
ATC CTG GAA CGG GTG GAT GGC CGG CGA CTC GTC TAC AAG TTT GGC AAA AAC TCA AGC GGC  
ile leu glü arg val asp gly arg arg leu val tyr lys phe gly lys asn ser ser gly  
1081/361  
TGG AAG GAG GAA GAG GTT CTC CAG AGT CGG AAC TGA  
trp lys glü glü glü val leu ser arg asn OPA

991/331 E8 E9

Fig. 4 (cont.)

11/21

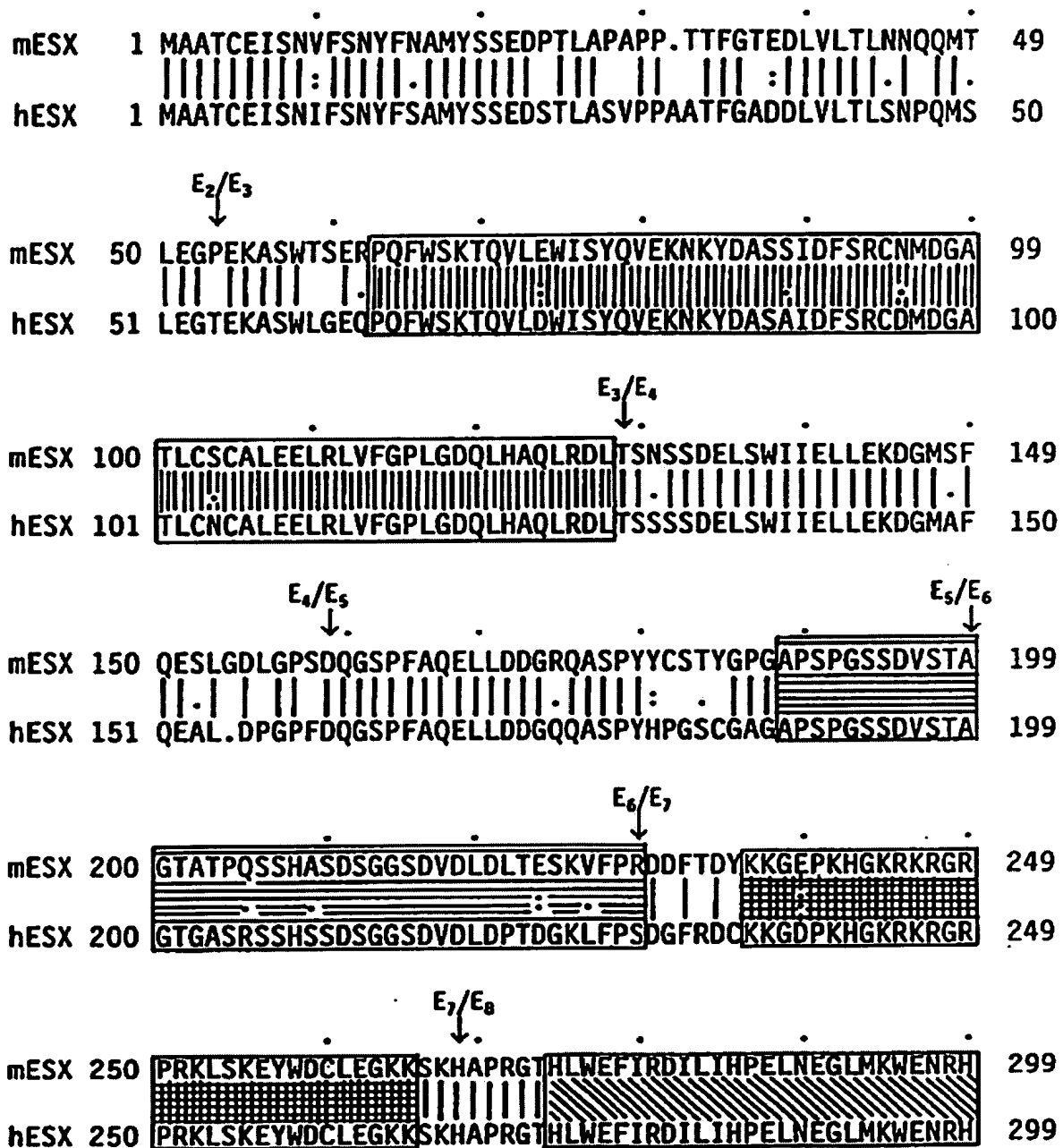


Fig. 5

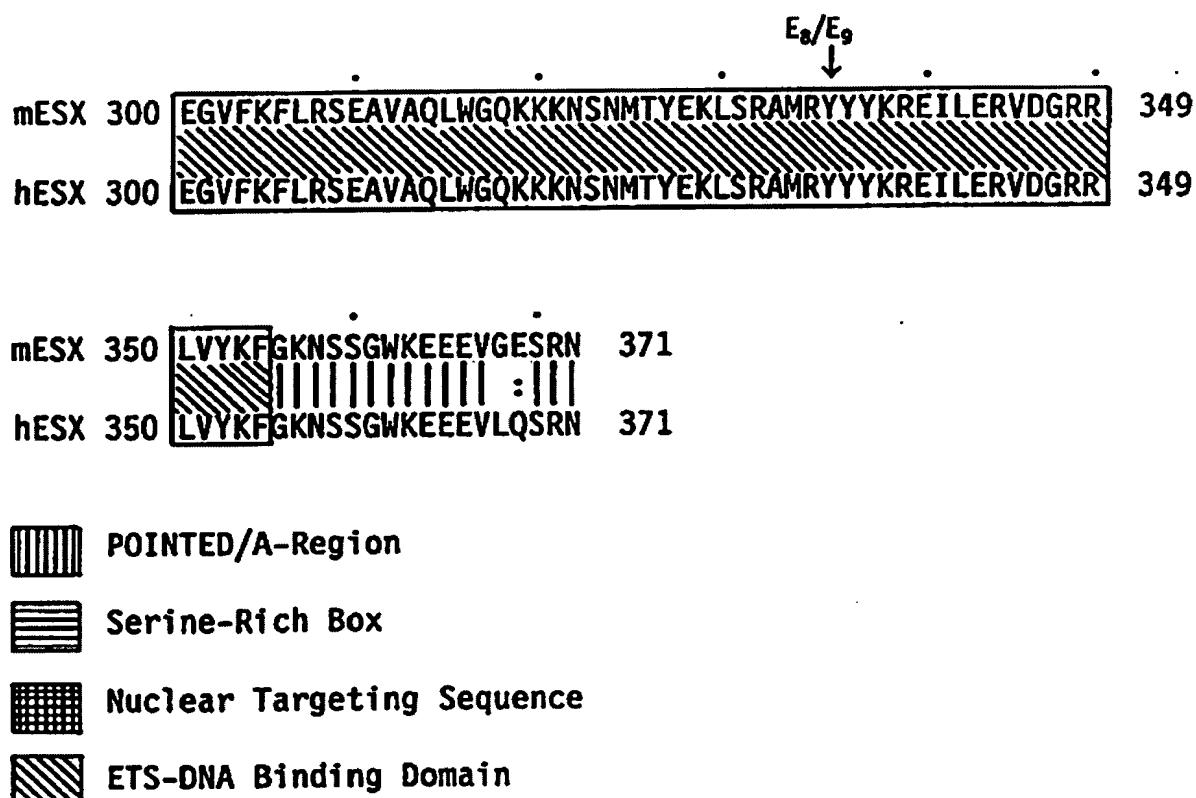


Fig. 5 (cont.)

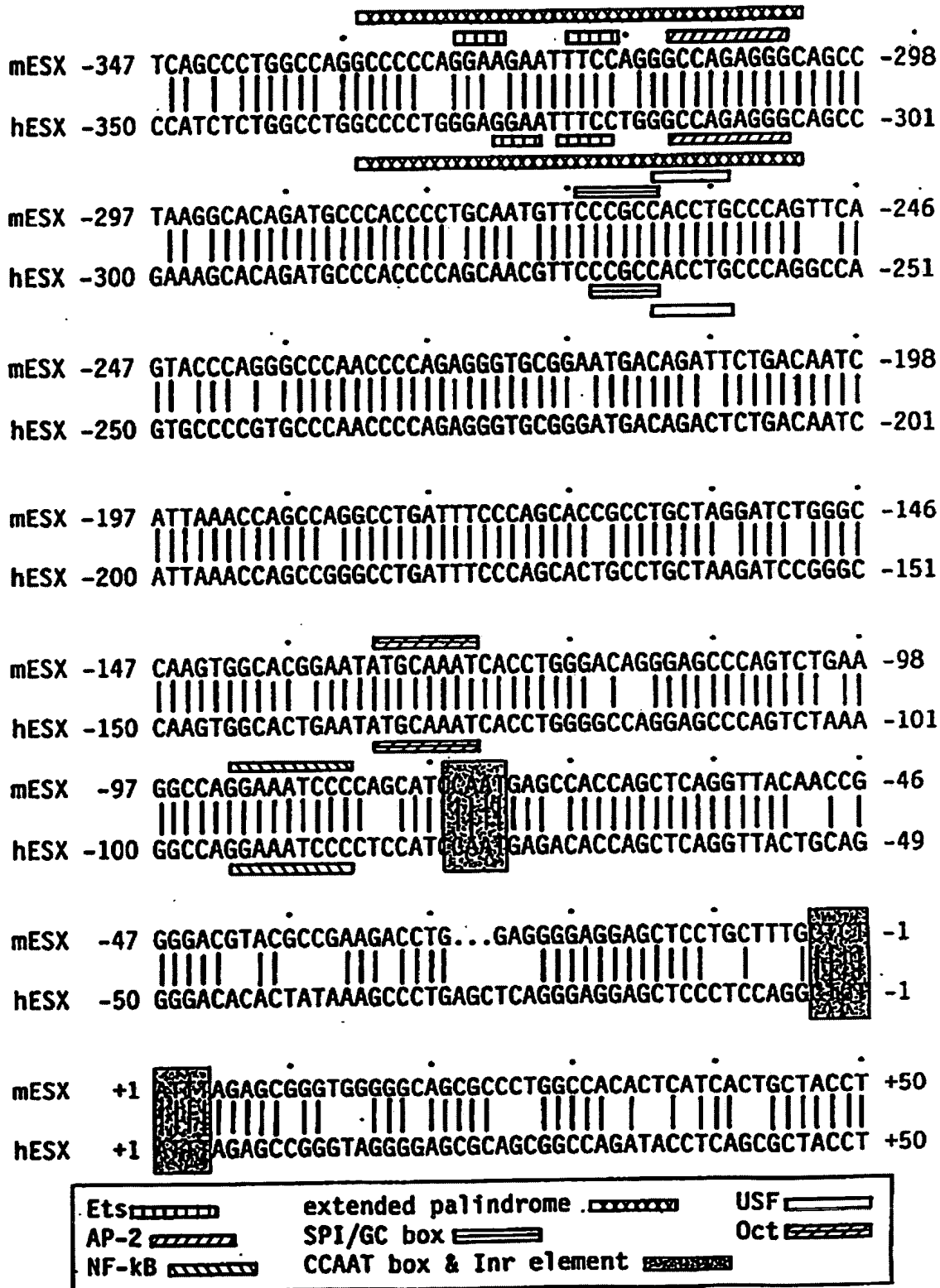


Fig. 6

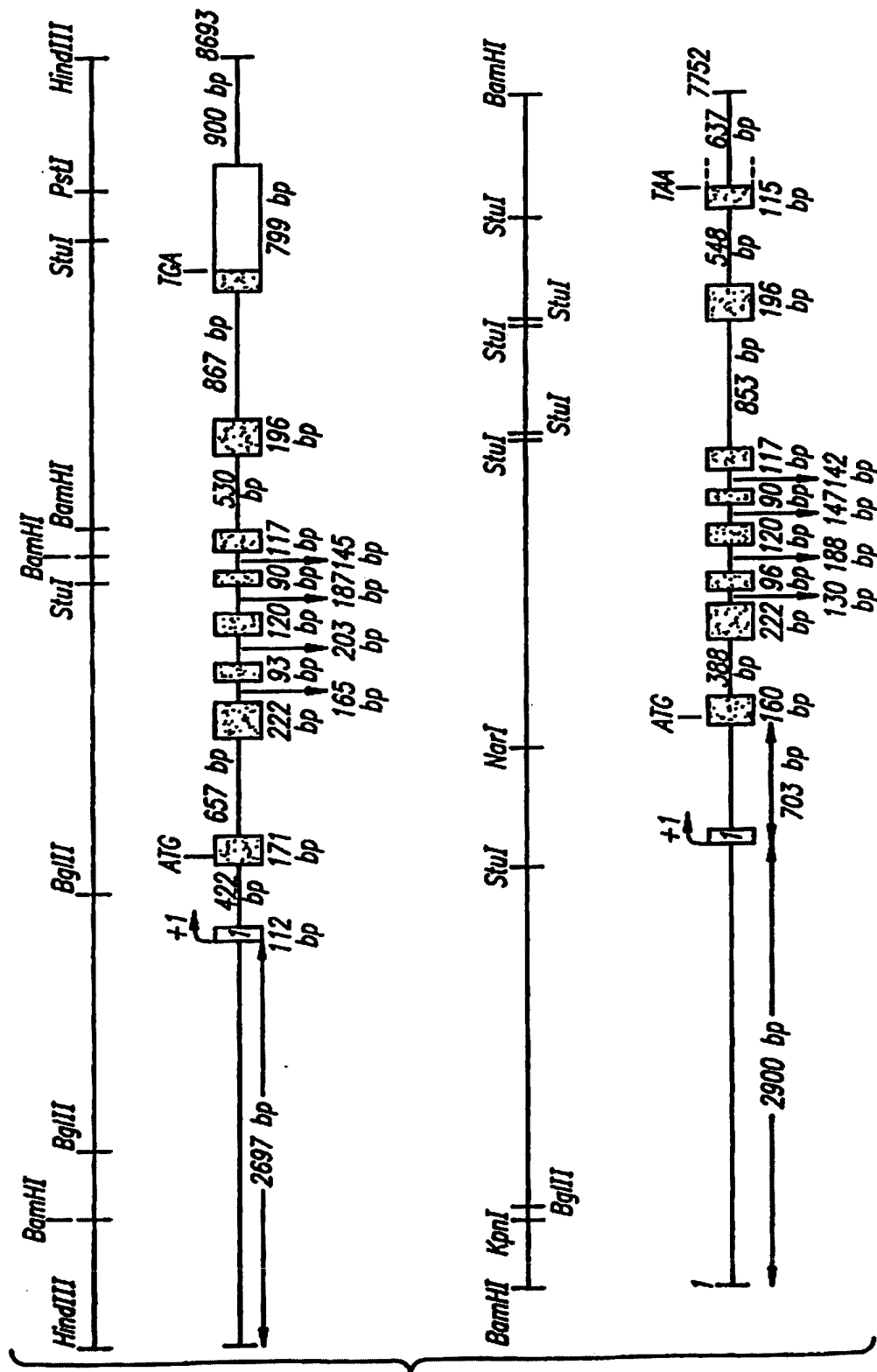
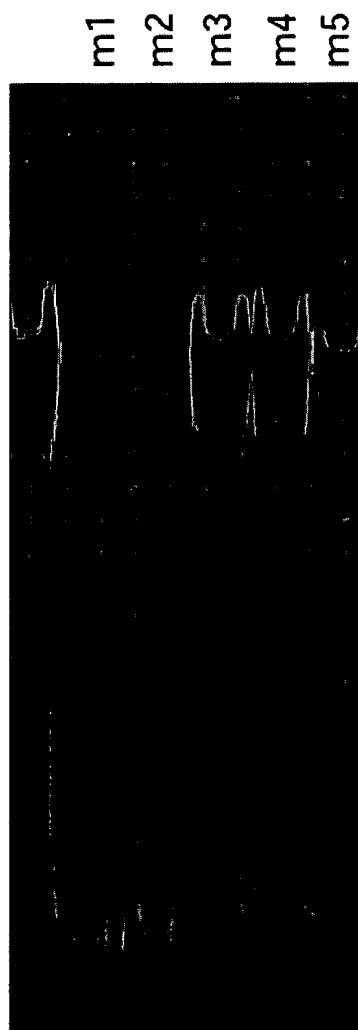


Fig. 7

15/21



WT 5' GGAGGAGGGCTGCTTGAGGAAGTATAAGAAT 3'  
m1 5' -----TA----- 3'  
m2 5' -----C----- 3'  
m3 5' -----AG----- 3'  
m4 5' -----CC----- 3'  
m5 5' -----C----- 3'

Fig. 8A

16/21

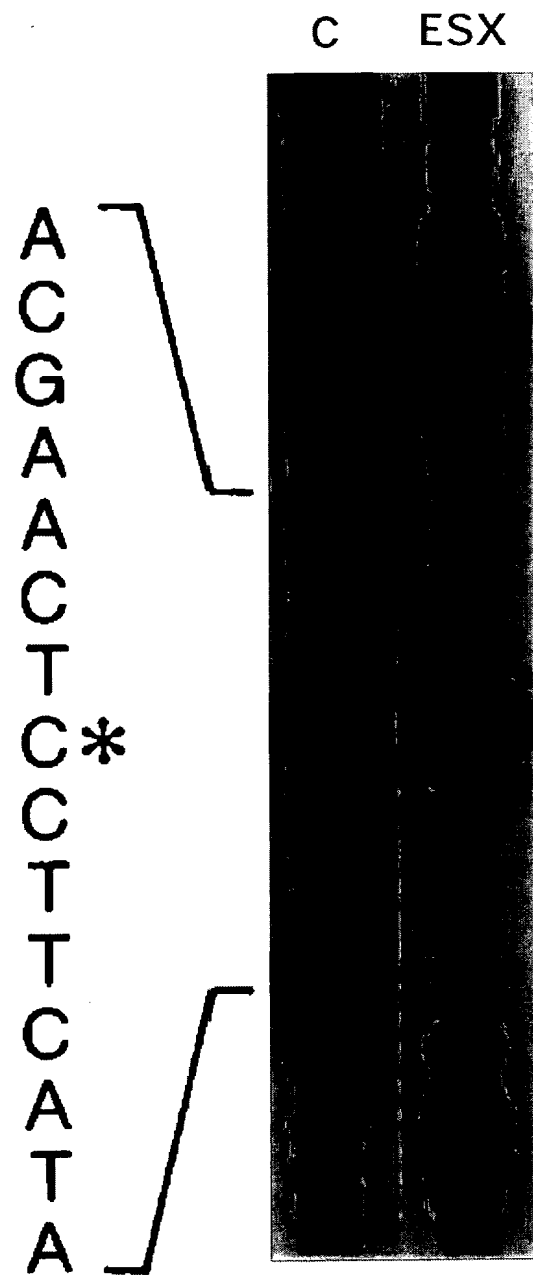


Fig. 8B



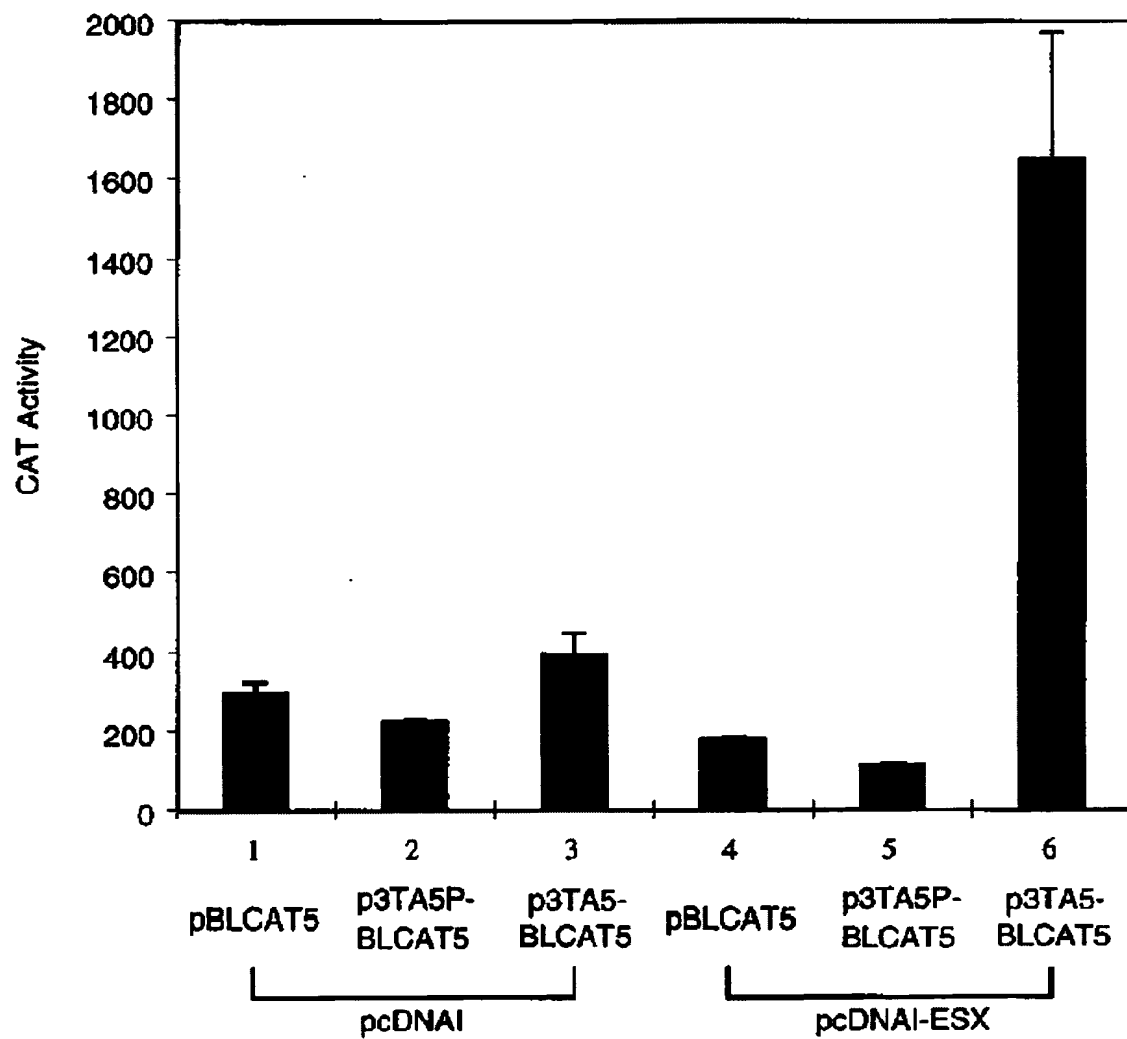


Fig. 8C

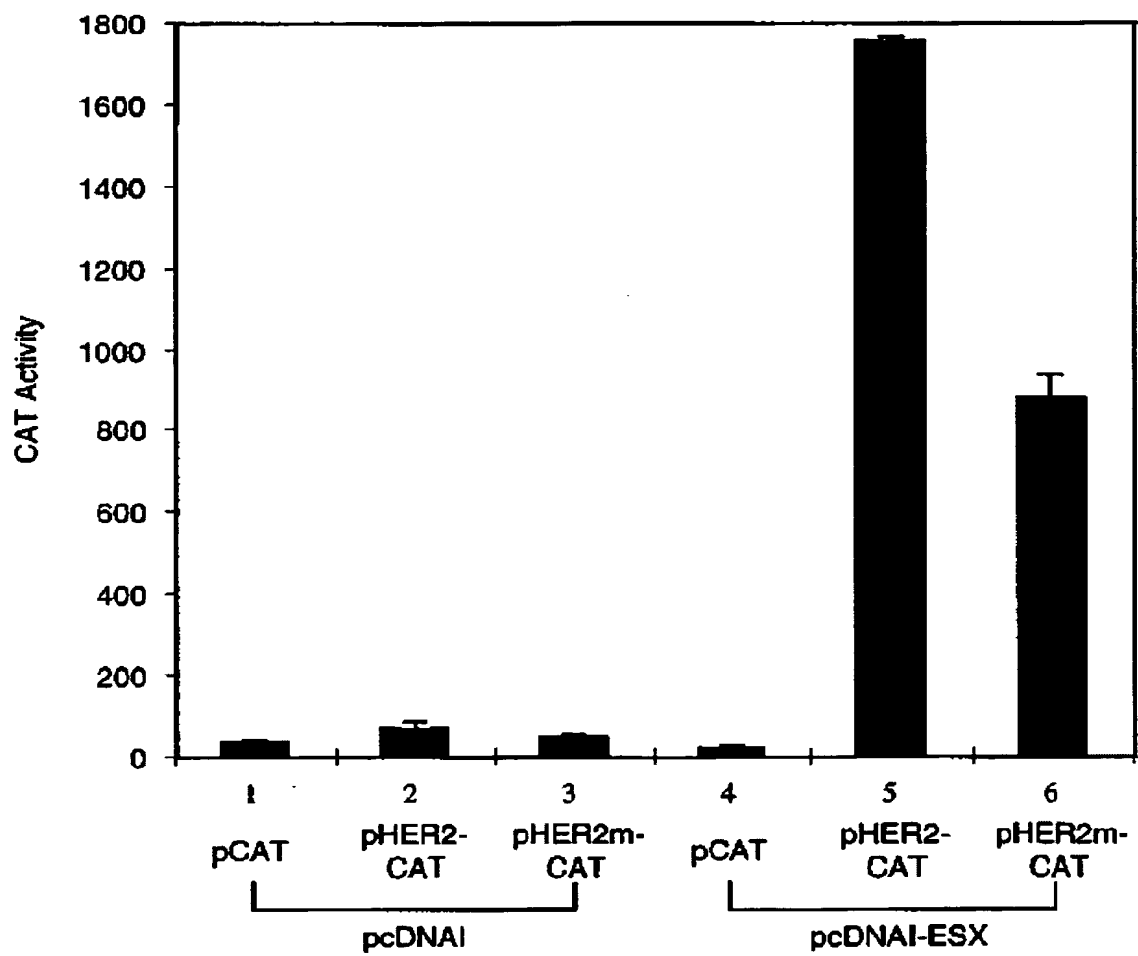


Fig. 8D

19/21

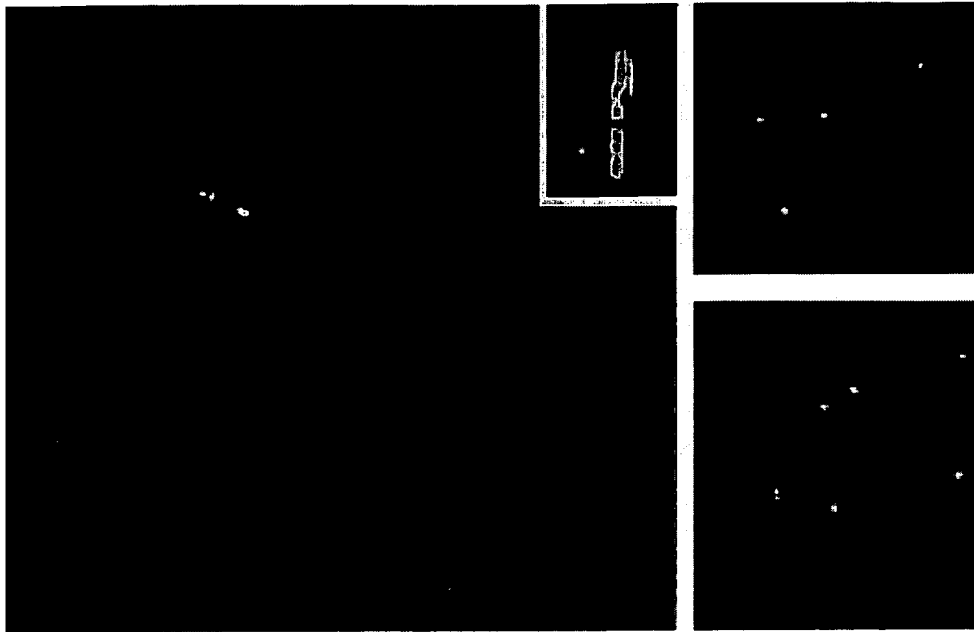


Fig. 8E

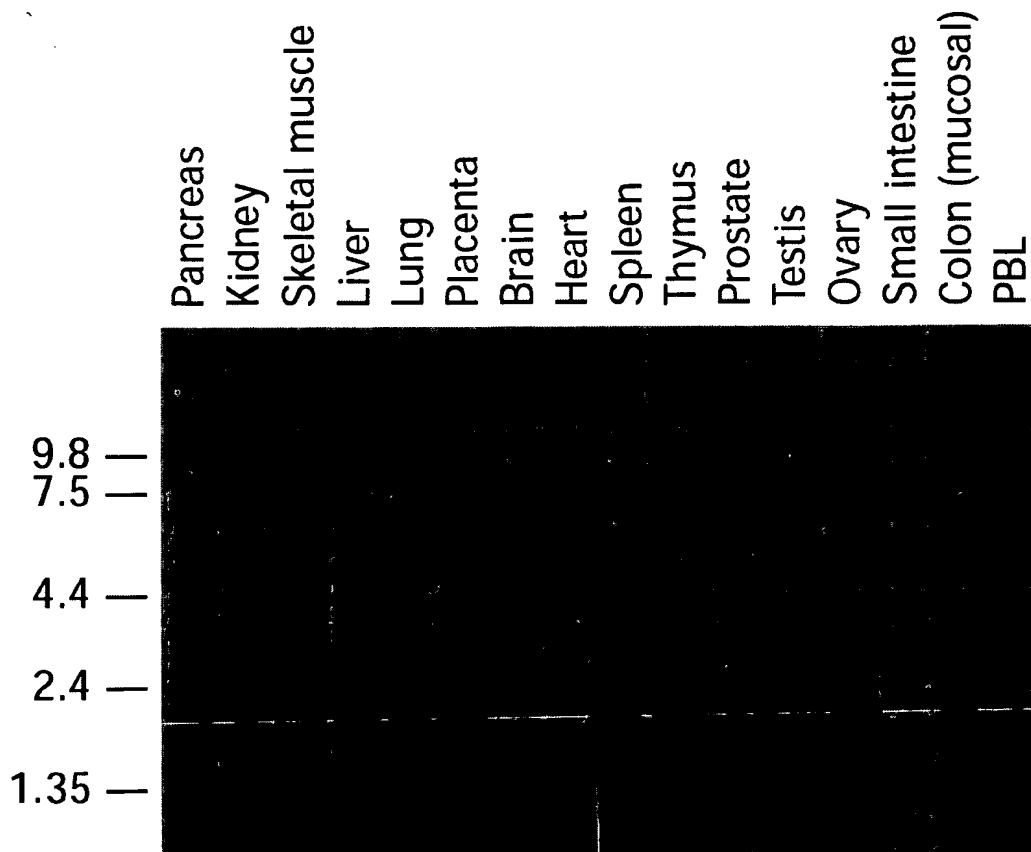


Fig. 9A

20/21

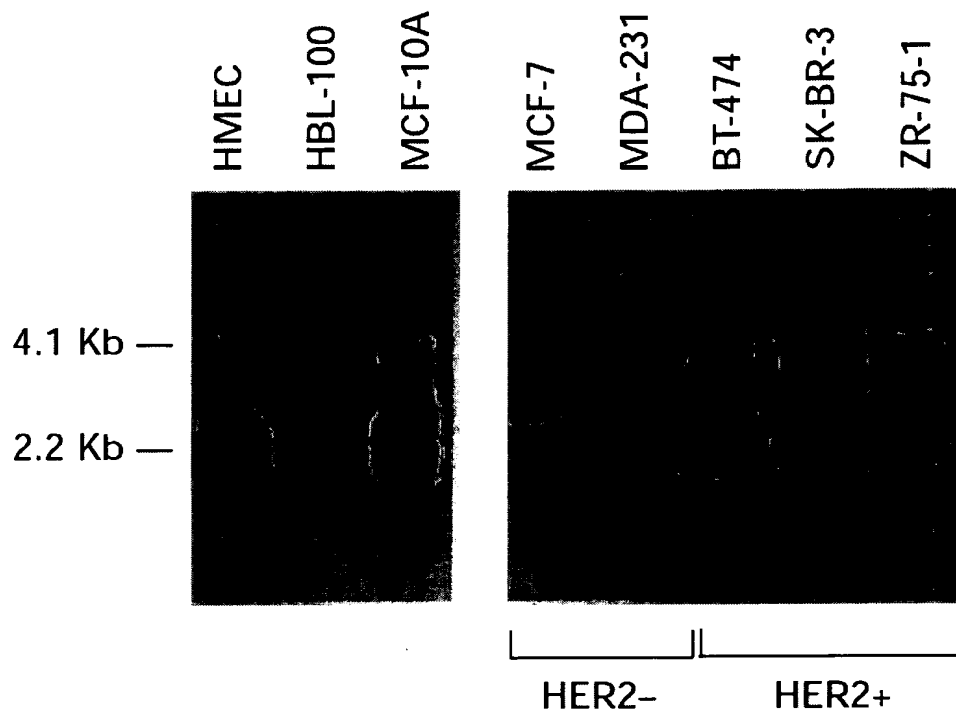


Fig. 9B

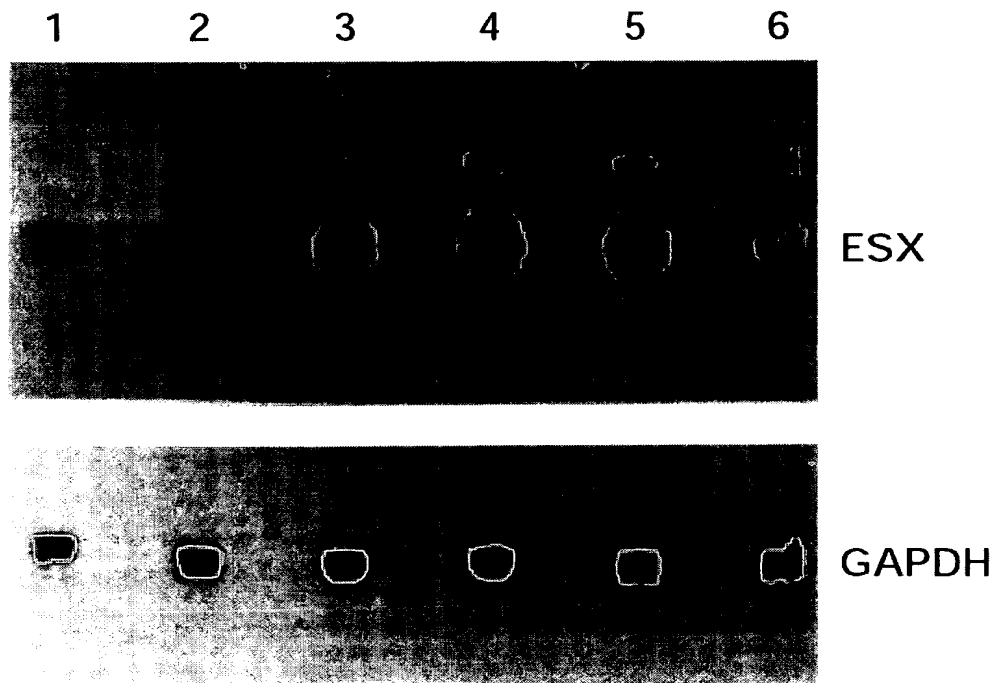


Fig. 9C

21/21

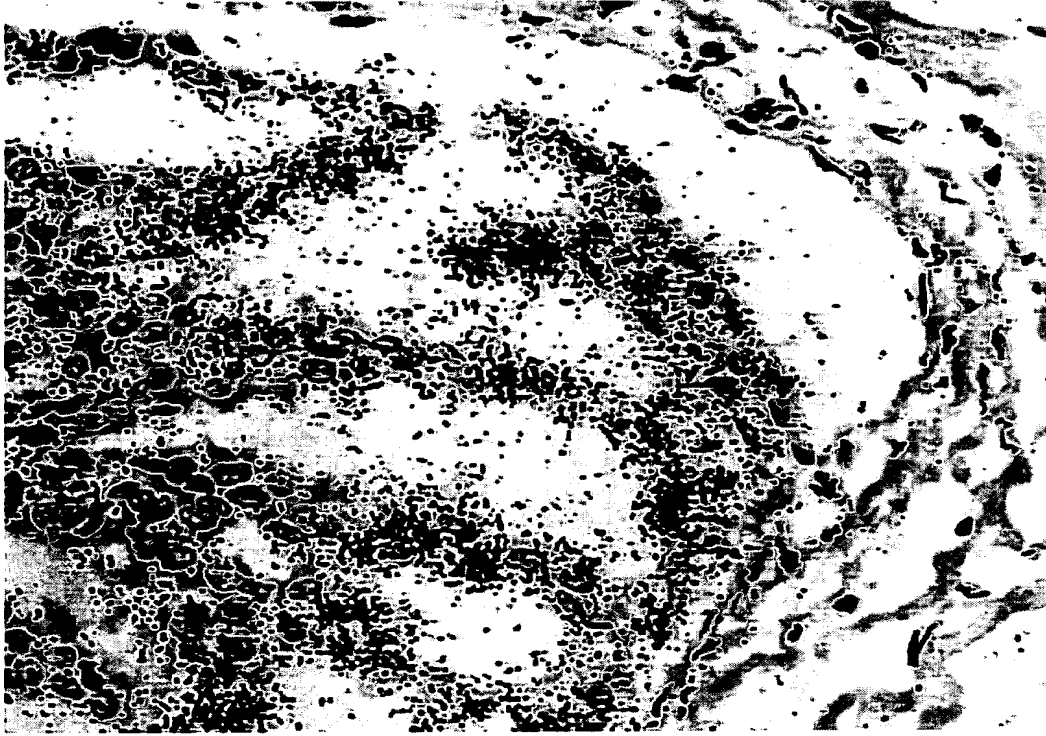


Fig. 10A

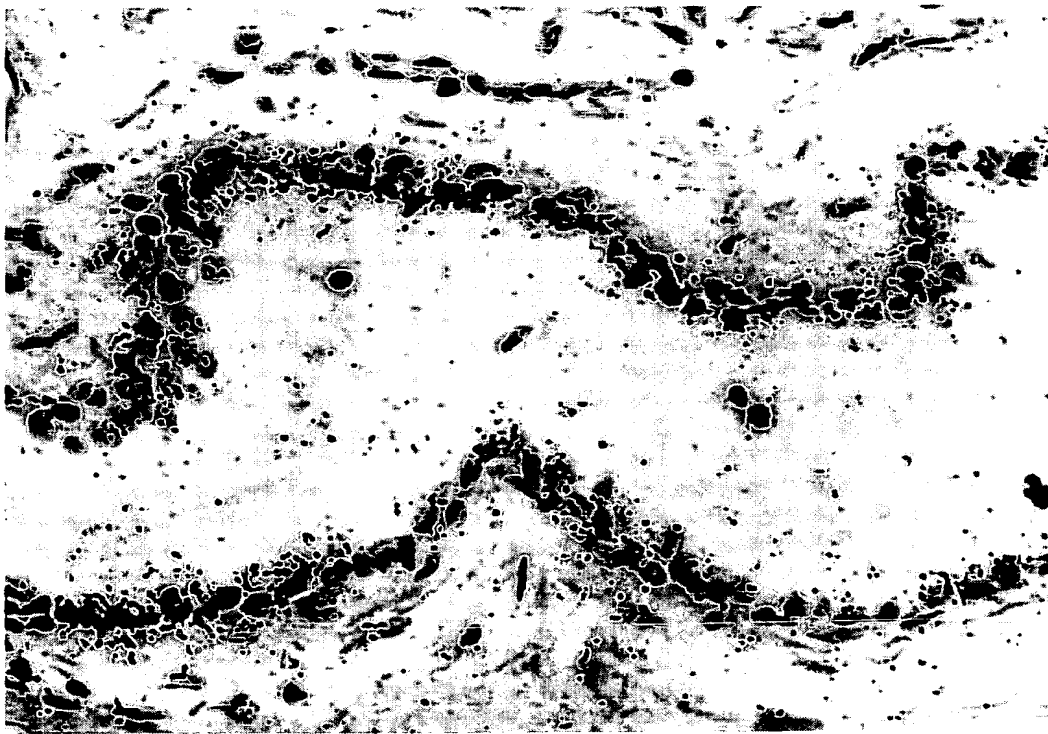


Fig. 10B